International Journal of Pharmaceutics 492 (2015) 316–331

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

The STEP database through the end-users eyes—USABILITY STUDY

Smita Salunke[1,*], Catherine Tuleu[1]

Department of Pharmaceutics and Centre for Paediatric Pharmacy Research, UCL School of Pharmacy, London, United Kingdom

ARTICLE INFO

Article history:
Received 21 May 2015
Received in revised form 12 June 2015
Accepted 15 June 2015
Available online 24 June 2015

Keywords:
Excipients
STEP database
Safety
Toxicity
Adverse effects
Paediatrics
Children
Neonates
Additives
Information resource

ABSTRACT

Objectives: The user-designed database of Safety and Toxicity of Excipients for Paediatrics ("STEP") is created to address the shared need of drug development community to access the relevant information of excipients effortlessly. Usability testing was performed to validate if the database satisfies the need of the end-users.
Method: Evaluation framework was developed to assess the usability. The participants performed scenario based tasks and provided feedback and post-session usability ratings. Failure Mode Effect Analysis (FMEA) was performed to prioritize the problems and improvements to the STEP database design and functionalities.
Result: The study revealed several design vulnerabilities. Tasks such as limiting the results, running complex queries, location of data and registering to access the database were challenging. The three critical attributes identified to have impact on the usability of the STEP database included (1) content and presentation (2) the navigation and search features (3) potential end-users.
Conclusion: Evaluation framework proved to be an effective method for evaluating database effectiveness and user satisfaction. This study provides strong initial support for the usability of the STEP database. Recommendations would be incorporated into the refinement of the database to improve its usability and increase user participation towards the advancement of the database.

1. Introduction

The acceptability and safety of excipients in relation to the age and development status of the child is the one of the central topic of discussion within the paediatric drug development community and health authorities. The choice and quantity of excipients has to be properly justified in terms of safety for the targeted paediatric population in Paediatric Investigation Plan (PIP) (Quijano Ruiz et al., 2014; EMA, 2012). The information and alerts of using an excipient or quantity of an excipient to be used in paediatric populations would provide the best approach to effectively mitigate adverse impacts of excipient by either avoiding the excipient or reducing the level of excipient in the formulation.

The database of Safety and Toxicity of Excipients for Paediatrics ("STEP") is created by European (Eu)[1] and United States (US)[2] Paediatric Formulation Initiatives (PFIs) to address the shared need of the paediatric drug development community to effortlessly access the available safety and toxicity information of excipients. The existing public sources (eg. TOXNET, Vitic database, Symyx, AcTOR etc.) do not provide complete and/or comparative information on safe use and acceptability of excipients in paediatrics and are built for different purpose (Fonger et al., 2014; Powles-Glover and Edwards, 2011). Also most of the existing databases have organized the information in free text format and

* Corresponding author at: Department of Pharmaceutics and Centre for Paediatric Pharmacy Research (CPPR), UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, United Kingdom. Fax: +44 2077535942.
E-mail address: S.Salunke@ucl.ac.uk (S. Salunke).

[1] On behalf of the European Paediatric Formulation Initiative (EuPFI).

[1] The European Paediatric Formulation Initiative (EuPFI) is a consortium founded in 2007 and working in a pre-competitive way on paediatric drug formulations. Members are from academia, hospital pharmacies, pharmaceutical industry (Innovators, Generics, Contract Research Organizations (CRO), Specials and Excipient Manufacturers) with European Medicine Agency (EMA) as an observer. The main objective of the members is to resolve scientific, regulatory and technological issues associated with paediatric formulation development.

[2] The United States Paediatric Formulation Initiative (US-PFI) is a project of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The PFI was established in 2005 to address the issue of the lack of appropri-ate formulations in children and to use this activity as a means to improve paediatric formulations.

http://dx.doi.org/10.1016/j.ijpharm.2015.06.016
0378-5173/

thus often do not have the capability to filter the data as per needs. A centralized information source that facilitates the analysis of large trends and simplified review on excipients use and acceptability in paediatrics is lacking and the STEP database aims to be the centre of efforts to address this gap.

The main objectives of the STEP database project are to: (1) identify, evaluate, compile, and integrate all available toxicity data on excipients, and (2) generate an interactive tool that can assist the paediatric drug development community in their assessments and management decisions regarding the use of excipients in paediatrics. The ultimate goal of this effort is the extraction/integration of information in a structured and easy to retrieve format that allows for on-the-fly data searches and specific searches based on the specific needs of the end-user to assist with the risk assessments of excipients and improve scientific decision-making through rapid access to centralized toxicity data.

As a part of the STEP database development process, the structure requirements and data elements were captured from potential end-users through global survey Salunke et al. (2012). Accordingly, an innovative user designed application is created to allow the end-user to navigate through a series of screens to select the desired data, narrow the information as per the type of data (clinical, non clinical in vitro, regulatory and reviews) and limit the results further as per specific needs (age, study dosage form, route of administration etc.) Salunke et al. (2013).

A beta version of the STEP database was released in September 2013 and was comprehensively tested both internally and externally by a number of individuals (testers). This step was critical to ensure that the application met the requirements expected of an interactive interface. During the testing window, recommendations were made to improve the appearance and functionality of the application, and a series of glitches and technical issues were identified and addressed. These early evaluations allowed the project team to address any outstanding issues and improve the presentation of the application prior to its public release. The end-result was the release of the STEP database Version 1 in September 2014 for systemic evaluation of its completeness, quality, configurability, usability and maintainability in real-world conditions.

The aim and objectives of this study were:

1. To validate the STEP database Version 1 (end-product) against the potential end users needs i.e. to insure the STEP database meet users' expectations.
2. To assess the database application functionality and usability by
   a. Ensuring appropriate ease of use (navigation), comprehension and user satisfaction.
   b. Characterizing how easy is to carry out a task using the database.
   c. Identifying problems in interacting with systems.
3. To assess the impact of this newly developed database on the paediatric drug development.
4. To set actionable recommendations to further improve the functionality of the system and increase its beneficial effects on paediatric drug development.

2. Method

Evaluation framework was developed to systematically assess the usability of the STEP database. Fig. 1 presents the framework used for the evaluation of the STEP database. It was adapted from the phases described by Kushniruk and Patel (2004).

2.1. STEP 1: Selection of testing approach

In general, usability testing is the process to assess the prototype in iterations, with changes implemented as necessary following each iteration with the goal of improving the product. Different usability-testing methods are used to assess health technologies and much literature exists that compares these methods. (Jaspers, 2009; Yen and Bakken, 2012). The methods commonly used are user based or expert based testing. This study implemented a user-based approach to assess the usability of a newly developed STEP database. The requirements on the type of information and structure of the database needed were captured from the potential user through global survey. Hence it was more appropriate to validate with potential users if the database developed met their requirements. Clearly, the ultimate test is the behaviour of real users interacting under normal working conditions but the user based testing method is an attempt to predict some or all of the issues that will occur in real use.

2.2. STEP 2: Sample selection and study design

The users who had provided their input on the data elements/structure of the database and had expressed interest in further involvement in the STEP database development process i.e. testing of the database as part of need assessment analysis were postulated to be suitable testers for the database. Being part of need assessment analysis, they were aware of the objectives and the concept of the database. They were approached and asked for their willingness to participate in the testing of the STEP database. Additionally few other potential end-users were also invited to test the database. These end-users were not involved in the requirement analysis but had expressed interest in testing the database.

The invited testers included regulators from European Medicines Agency (EMA), members (majorly toxicologists) from preclinical paediatric toxicology group led by Janssen, members (majorly formulation scientists) from Innovation and Quality (IQ) consortium, academic professionals and PhD students. No limit was set on the number of the users that could participate in the testing. There has been much debate about how many participants are needed in a usability test to reliably identify usability issues. Prior studies have shown that carefully conducted usability studies involving as few as 5–8 subjects can lead to identification of up to 80% of the surface level usability problems with an information system (Friedman et al., 2006; Nielsen, 1993). For STEP database it was anticipated that at least 5 subjects from each group would complete the test.

Simple study design was used. The group of participants identified above interacted with the database (with each participant carrying out the same task or set of tasks) in order to assess problems with the design and functionality of the STEP database. However for future testing, within-group designs would be utilized, where individuals may be asked to try out revised versions of the database, or participants from this study may be followed over time as they learn how to use the database.

2.3. Step 3: Selection of test method

A wide variety of approaches and methodologies have been applied in assessing the impact of information systems in health care, ranging from controlled clinical trials to use of questionnaires and interviews with users (Levi and Conrad, 1998; Tullis, 1998). However, application of these methods varies according to the system being tested and resources required to use them effectively.

For the STEP database evaluation, a hybrid of different methods was used. The test consisted of three parts: demographic questionnaire, scenario based tasks, and a satisfaction

questionnaire. Empirical testing of the database by giving a scenario or situation to users and asking to perform the task by using the database, not only enables the testing of the usage of the database but also assesses the user's satisfaction in terms of the design and structure of the database (Vrazalic, 2003). Hence a *scenario-based usability test* was the best-fit method to validate if the users requirements have been met. It involved representative end-users and scenarios or specific tasks designed to cover the major database's functionality and to simulate expected real-life usage patterns.

In addition to scenario based tasks, questionnaire was designed to assess the less observable aspects of interface design that

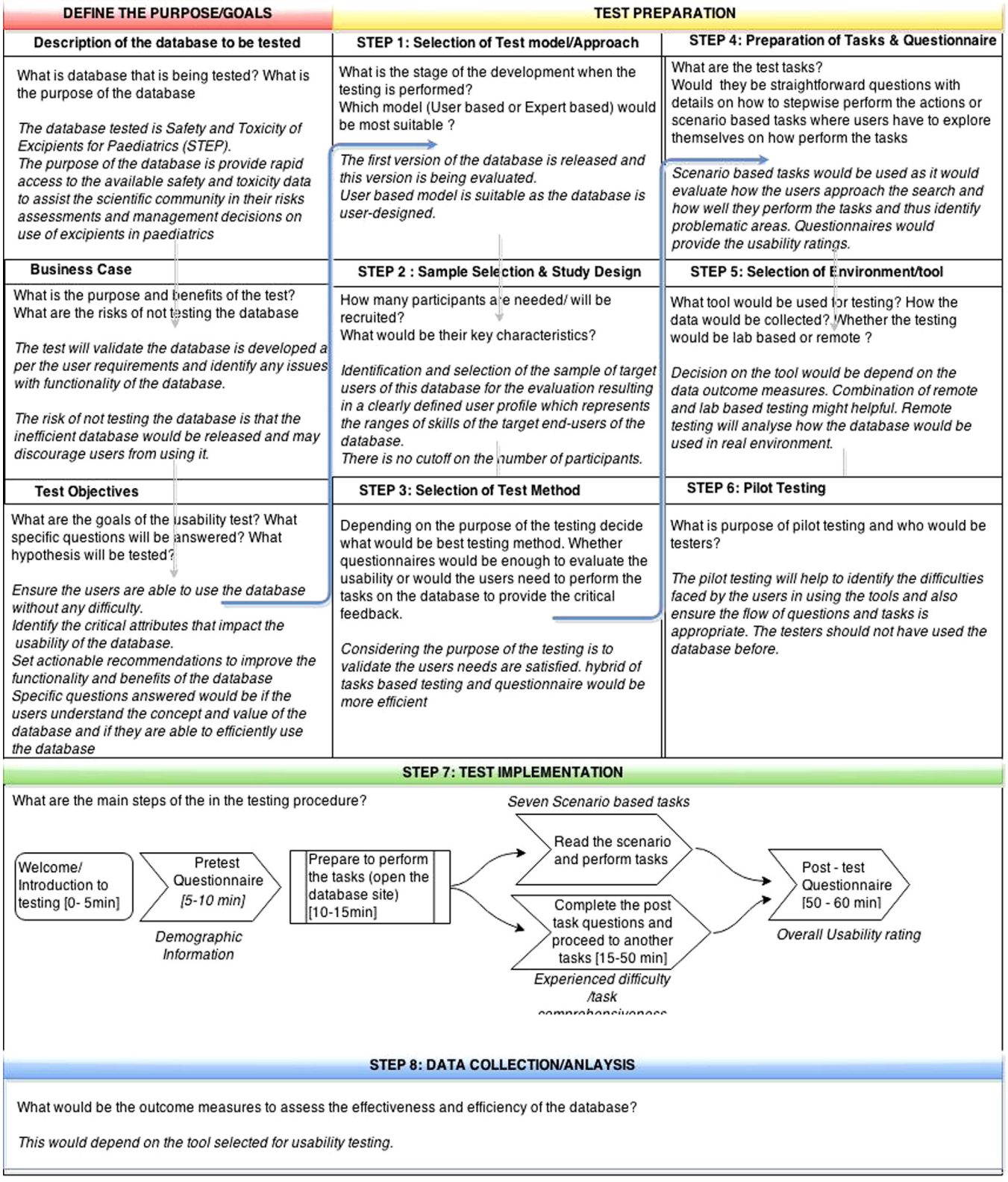

Fig. 1. The STEP database usability evaluation framework.

cumulatively contribute to a user's subjective feelings of satisfaction or frustration.

2.4. Step 4: Preparation of tasks and questionnaires

2.4.1. Scenario based tasks

The test consisted of seven scenarios and each scenario tested one or more tasks. Each of these tasks, were directly linked to the goal of the study (Appendix A). The tasks addressed particular areas of concerns regarding navigation or understanding of database structure or data fields. There were two major aspects to be considered while preparing the scenarios and tasks for the STEP database; the user job role (e.g. formulation, toxicologists, regulators) and different functionalities provided by the database. The database would be used by wide range of professionals (formulation scientists, toxicologists) who have different responsibilities and therefore would use different features provided by the STEP database; for instance a formulation scientist formulating a dermal product for children may use the database to find out which excipients have been used dermally in children and its safety data. The toxicologist on other hand might look for species used in specific toxicity study (e.g. species used for carcinogenicity studies). Hence to get a better insight of the opinion on database interface and usability of the database, the scenarios were developed according to the role of the end-users, while the tasks were designed in a way that would assess the effectiveness and ease of use of the database. The tasks that would be commonly performed by the wide range of end-users were tested.

- Log in.
- Use of single field (Search for a specific excipient).
- Use of multiple fields (Search for excipients by age and route of administration).
- Use of filters for refining the results.
- Ability to view all records from a search at once.
- Ability to broaden a search.
- Ability to narrow a search.
- Ability to export the data.

As part of post-task questionnaire, participants were asked to rate difficulty of task completion immediately after completing each scenario/task using a five-point ordinal scale with 1 indicating very easy and 5 indicating very difficult.

2.4.2. Questionnaires

Questionnaires have long been successfully used to evaluate user interfaces (Root and Draper, 1983). The demographic (pre-test) and satisfaction (post-test) questionnaire was designed using Qualtrics online software (Appendix A). The purpose of the pre-test questionnaire was to obtain demographic information that would give a better understanding of the organization and job function of the end-users and the extent of their experience with online searching. As the database is newly developed and not accessible to end-users, questions on familiarity or pre-use of the database were out of scope.

The satisfaction questionnaire was developed to assess the overall satisfaction level of users. It supported testing and determination of user subjective satisfaction with the database interface, as well as user satisfaction with its ease of use, efficiency, likeability, as well as the attitude towards the database induced in users during its usage. The questions mainly involved rating the specific element of the database usability with few open ended questions as in; what they particularly liked or not liked about the database and their opinion on the ways on how the database can be improved.

2.5. Step 5: Selection of environment/tool for testing

2.5.1. Selection of environment

The next step was the selection of environment and tool for testing. The physical location of the evaluation can vary considerably depending on the type of application and usability attributes. Two major settings commonly used for usability testing of applications are (1) laboratory experimental studies (2) field/

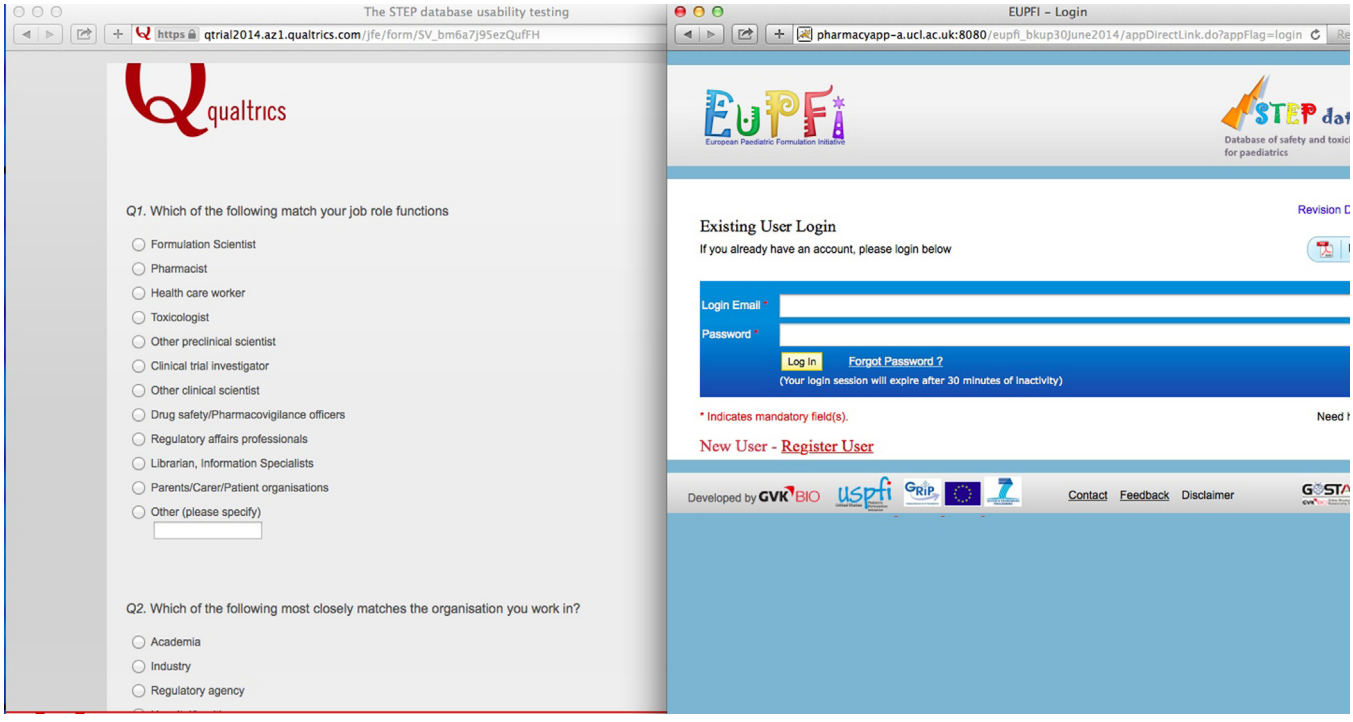

Fig. 2. Qualtrics tool.

remote usability studies (Kaikkonen et al., 2005). Remote testing environment was deployed for STEP database to enable participants to provide feedback based on their experience with the database in a read world setting (Power et al., 2009).

2.5.2. Selection of tool

Remote testing the database needs an external software tool or an instrument to present the scenario-based tasks to the participants and then to capture their behaviour or response. Several software tools are being developed and available on Internet to increase the efficiency of user tests. The usability testing tools explored for the STEP database testing were Usabilla, Usabilitytools, UsabilityHub, LOOP, User Testing, Zoom and Testrockit. The minimum requirement was that the tool should be cost effective, allow more than 20 users to test the database at a time, user friendly, be able to perform scenario based testing and to collect the user feedback on difficulties faced in performing the tasks.

Testrockit was selected, as it met all the basic requirements for the STEP database usability testing. It is free/open source tool for usability benchmark studies and delivers usability metrics for success, efficiency, and satisfaction in compliance with the ANSI/INCITS-354Common Industry Format (CIF) for Usability Test Reports.

In TestRockit, usability studies are delivered through a web-based widget that presents the user tasks and collects their input (West and Lehman, 2006). The major limitation of Testrockit tool was that it required downloading of Microsoft Silverlight software by the participants. Hence, another method of testing was explored considering some participants may not be able to download the software as they use their work computers. Also not all users are very proficient with computers and may be not be aware how to download or use new software. A simple questionnaire based testing was opted, since most of the users are generally familiar with questionnaire-based surveys. Qualtrics online survey software was selected to create the questionnaire (Appendix A) because the same platform was used to design to pre and post-test questionnaires. The drawback of using Qualtrics was that it is a survey tool and not a usability testing tool and hence required more initial set-up for the actual assessment. The participants would have to manually access the database and switch between two webpages (survey and database) as they completed the eight scenarios/tasks" (Fig. 2). While in the Testrockit, the database opened automatically as study was launched and was by-default visible on the same webpage as the question/task panel (Fig. 3). The URL links to both the tools were provided to the participants and given a choice to select either tool as per their preference or computer efficiency.

2.6. Step 6: Pilot testing

The pilot testing was performed with 10 EuPFI members representing wide range of professions from scientists to regulators. The EuPFI members were first given an introduction to the database and explained how to perform the testing. Some members used the Testrockit tool and others used the Qualtrics survey. The pilot testing helped identify the difficulties faced by the participants in using the tools, made sure the database worked correctly, as well as tested if the flow of the questions were in appropriate order and would not create any problems when conducting the test.

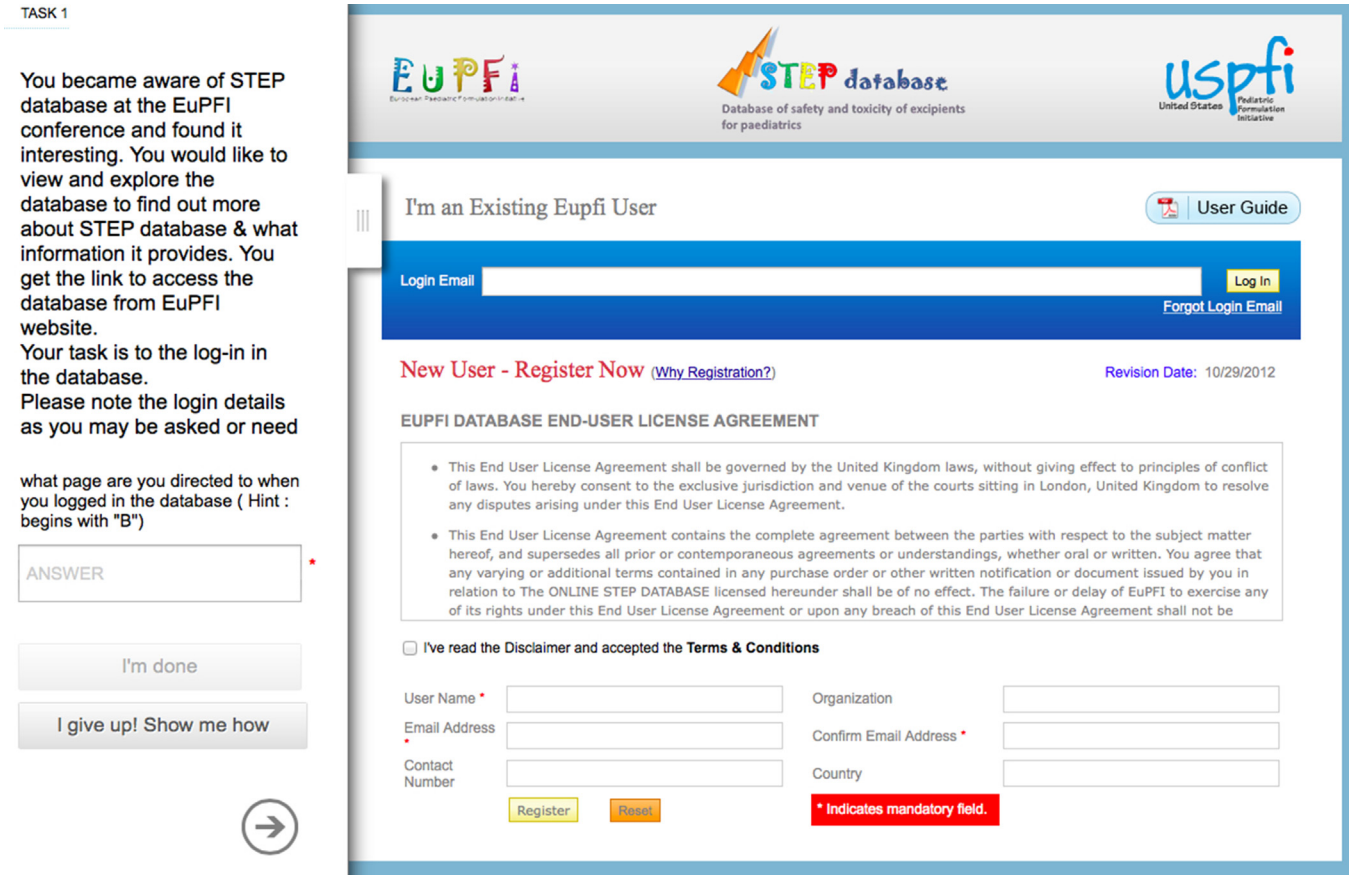

Fig. 3. Testrockit tool.

2.7. Step 7: Test implementation

A series of usability tests were carried out with different groups as described in Step 5. In general, the process involved three steps (1) a pre-test survey to collect the data about participant background and computer and Internet experience (2) a scenario based testing which required participants to use the database to find the specific information about the excipients and their safety and toxicity (3) a post-test survey which consisted questions and statements about their satisfaction on perceived usefulness and ease of use, as well as the navigation, content and appearance of the database.

For testing that was performed locally [regulators at European Medicines Agency (EMA)], one-day workshop was organised and the facilitator was available to prompt the participants and deal with any technical issues. While for remote participants (IQ Consortium members, Preclinical paediatric group and Novartis paediatric group) a pre-and post webinar was organised. In the pre-webinar session, the STEP database and testing tools (Testrockit/Qualtrics) were introduced and demonstrated to the participants. An invitation letter with hyperlink to the testing site was sent to the webinar participants. The letter explained the purpose of the study and encouraged users to participate in the testing to help improve the database. The participants were given two to three weeks to perform and complete the testing. A post webinar was organised two–three weeks after the pre-webinar to receive the feedback on the database and discuss on the problems and shortcomings of the actual testing process.

In addition to above testing, a workshop was organised at the 8th AAPS Italian University Network (A.It.U.N.) Annual Meeting on "Medicines for Children's safe: challenges and opportunities" on March 6–7th, 2014 at University of Pavia (AAPS, 2014) to test the database. The aim of this meeting was to discuss the problems related to the paediatric medicines from different points of view: the hospital pharmacy perspective with the need of everyday extemporary formulations and parenteral nutrition, the clinical aspects and the necessity of planning suitable clinical trials to collect enough data, and the formulation issues focused on preparing safe and compliant medicines. The meeting provided the appropriate representatives for testing the database. The database testing workshop was attended majorly by PhD students and professors involved in paediatric formulation research. Other participants were from Italian regulatory agency and pharmaceutical industry. Only paper-based Qualtrics survey was provided to the participants to provide their feedback.

2.8. Step 8: Data collection/outcome measures

The fundamental purpose of the usability evaluation was to quantitatively and qualitatively measure the effectiveness and efficiency of the STEP database and as well as the user satisfaction as they performed the series of tasks. As the test involved using two different tools, the output measures varied. TestRockit tool had inbuilt ability to collect the information on: time on task, task success and failure, task quitting (Instructions were provided to complete the task and users identify the step that caused them to fail), task satisfaction and overall product satisfaction (basic rating or Net Promoter Score). While for Qualtrics all the above output measures had to be calculated manually on basis of the answers provided by participants. The feedback received for the user satisfaction was coded to get the information on user satisfaction. The Qualtrics survey gave the start and end time for each completed survey but not for each task. Qualtrics did not provide the "Net promoter score" score but that was not the mandatory requirement to assess the usability.

Percent task success rate was used as the key performance metrics for this study as this measure could be obtained with the both the tools. Percent task success rate measured the percentage of a given task that participants successfully complete without critical errors.

2.8.1. Database usability/user satisfaction

In addition to the percent task success rate, data on user satisfaction (*attitude towards the database after using it*) and feedback on the database in terms for further refinement were collected in several ways: response to a series of questions following each task and to the post-test questionnaire. The post-test questionnaire included simple ten-questions that provided the comprehensive assessment of usability. The questionnaire was completed by each participant after completion of seven scenario based tasks, but before any debriefing took place. Participants responded to five-point Likert-scaled questions on information content (e.g. whether the information system provides too much information or tool little), Comprehensiveness of presentation (e.g. the way the database is structured as in different sections for different type of data) problems in navigation or ease of use of search interface and overall system understandability and impression of the database.

3. Data analysis

Analysis focused on two main areas:

1. The usefulness of the STEP database in terms of its content. This mainly dealt with the issues such as whether the STEP database provides useful, up-to-date and valuable information to a user. The responses collected were also analysed to assess the data quality in terms of data accuracy, data integrity data completeness and data availability.
2. The ease of use of the database or interface. This characterized the potential problems or issues related to the actual user interface or database design and structure.

4. Results

4.1. Usability evaluation response rate

Overall 66 potential end-users participated in the usability evaluation of the STEP database. However, not all participants answered every task or question. Some participants only completed the scenario-based tasks and did not complete the post-test questionnaire. While some participants only completed the post-test questionnaire and did not perform the scenario based tasks. To get a complete picture on how a user would use the database, difficulties faced during the use and attitude towards the database after using the database, it was necessary that participants performed all the three steps of the testing; pre-test; scenario based tasks and post-test. Hence only the responses from the participants who completed all three tests were considered further for usability analysis. Overall out of 66, 29 participants completed the usability test and provided suggestions for the STEP database improvements throughout the session, resulting in a response rate of 44%. According to usability experts, this was the sufficient sample size for the first round of usability testing (Faulkner, 2003; Virzi, 1992).

The four main categories of users for this study included toxicologists, regulators, formulation scientists and academic researchers. They are anticipated to be major users of the databases. The participants included 9 toxicologists, 10 regulators, 5 formulation scientists and 5 academic researchers and were

mostly from Europe and United States. All were adequately experienced with using a computer and the web.

4.2. Overall user performance: effectiveness and efficiency

Al the participants completed the same number (7) of tasks. Although the participants represented different professions such as toxicologists, regulators, academics and PhD students, there was no major difference in the performance of the tasks (Table 1). Neither the level of experience in drug development nor knowledge of formulation development affected performance. This indicates that the database is neutral and comprehensible to wide range of professions.

4.3. User performance with individual tasks

The performance for individual tasks is summarized in Table 2. The breakdown on participants' success, problems, comments and changes implemented as per each task is presented in Table 3. Overall, 41–100% of the participants completed all the tasks successfully. Apart from the comments summarized in the Table 3, the other key comments for improvements include:

1. On the "Search BY Excipient" page once excipient is selected all the other search categories (CAS number and Synonyms) self-populate with information specific to excipient except the "Function" category. The user requested for the function category to be self-populated with excipient-specific functions.
2. Displaying the total number of references as per the data type (e.g. number of references for clinical data or non clinical data).
3. Provide short definition of Search BY Excipients and Search For Excipients.

4.4. User satisfaction

As presented in Fig. 4, overall 93.1% of the participants had good general impression of the database. In terms of data content and structure of the database, 86.2% found the information content was good, at the sufficient level needed and the presentation of the results were easy to follow. While in terms of navigation and ease of use of the STEP database, only 38% of participants were satisfied with the functionalities available to the search or use the database and 52% were satisfied with the level of the support information available to help users.

4.5. Qualitative assessment

To gain the deeper understanding of the subjective aspects of the interaction between users and the STEP database, the qualitative approach was implemented with focus on user satisfaction (Pope and Mays, 1995). The open ended questions and the de-briefing sessions after the usability testing helped gain the feedback both on the technical failures as well as the user' attitudes to content, usefulness of the STEP database for paediatrics and acceptance of the database in scientific community. According to the discussion with participants in the debriefing sessions and answers to the open ended questions, the ultimate strength of the database was the information content. Overall users were very satisfied with the level and different types of information available for each excipient in the database. The comments on likes and dislikes of the database are summarized in Table 4.

Around 86% of users agreed that the database is important and useful and that they would use it for searching the information on excipients safety and toxicity as and when needed. 14% were not sure, as the database has very limited excipients (10 excipients during usability testing) and it was difficult for them to envision the possibilities and the applicability of such a system with limited information.

5. Discussion

For large scale web based scientific research applications usability testing is very important since the purpose of such application is to facilitate research, allow collaboration among its users, share personal experience, and accelerate the dissemination of knowledge. The purpose of the STEP database usability testing differed from the typical usability studies performed on information system. Unlike these usability studies, which are only targeted to assess the user interaction with the system, the STEP database usability study had an additional objective of validating if the system was developed as per the user requirement and that it achieves acceptable performance levels. The idea was to develop a user evaluation approach that can be used in the iterative evaluation of the STEP database during the development (i.e.

Table 1
User performance by experience.

Scenario	Post task questions	Number of participants who completed tasks successfully		
		Toxicologists (out of 9)	Regulators (out of 10)	Academics/formulation scientist (out of 10)
1. Log-in	Question 1	8	6	5
	Question 2	8	6	5
Search BY Excipients	Question 1	7	6	6
	Question 2	5	5	5
3: Results page—clinical data filtering	Question 1	7	8	9
	Question 2	3	5	4
4: Results page—non clinical data filtering	Question 1	9	9	9
	Question 2	7	7	7
5: Search FOR excipients	Question 1	5	6	6
6: Refine specific excipient using search for excipient	Question 1	8	8	9
	Question 2	7	6	6
7: Export the results	Question 1	9	10	10
	Question 2	9	10	10

Table 2
User performance with individual tasks.

Scenario	Post task questions	Number of participants who completed the task successfully
1. Log-in	Question 1	25(86%)
	Question 2	27 (93%)
2. Search by excipients	Question 1	26 (89%)
	Question 2	26 (89%)
3: Results page—clinical data filtering	Question 1	24 (83%)
	Question 2	12 (41%)
4: Results page—non clinical data filtering	Question 1	27 (93%)
	Question 2	21 (72%)
5: Search for excipients	Question 1	17(59%)
6: Refine specific excipient using search for excipient	Question 1	25(86%)
	Question 2	19(65%)
7: Export the results	Question 1	29 (100%)
	Question 2	29 (100%)

formative evaluation), with the objective of improving the design and deployment of the database as well as ensuring that the process of design of database leads to effective system.

A database is only useful for storing information if the users can correctly and efficiently query the database and the developers can update, and maintain the existing data to inform decisions or to perform assessments. Consequently, critical steps in assessing the effectiveness and usefulness of the STEP database dealt with users ability to perform the scenario based tasks, or real search situations, as well as system' ability to provide information searched by the user to help them make inform decisions based on the current state of knowledge. Testing consisted of performing data queries as per the scenarios given and selecting the appropriate filters with associated toxicological data, or user-specified query criteria.

The results of this study matched the findings of the previous research on benefits of using multi method approach for usability testing (Qiu et al., 2007). Hybrid method of scenario based tasks and open-ended questionnaires helped assess a variety of questions ranging from analysis of specific interface problem to overall assessment of user expectations. The scenario based tasks helped assess the direct experience of the user and identify the system-dependent attributes (information content, ease of use and user support) that have major impact on usefulness of the STEP database and user satisfaction. While the questionnaire and debriefing session identified independent attributes including personal attitudes of the users towards computer use as well as the attitude of the user towards using the STEP database. In addition the usability scores indicated the perceived usability of the database. The scenario based tasks helped determine how and when specific usability problems occurred, which would have been difficult to assess with questionnaire alone. Certain navigation problems and inadequate instructions were identified in both methods and hence were considered as critical attributes affecting the usability of the STEP database.

While the main issues and lesson learned about the database through the usability testing were valuable and some integrated into the refinement of the database, the three critical attributes that were identified to have impact on the usability of the STEP database included (1) content and presentation of the results in the database (2) the navigation and search features (3) potential end-users of the database.

5.1. Content and structure of the database

Perceived usefulness and learnability of the information system are key determinants of the end-user satisfaction (Calisir and Calisir, 2004). End-users are likely to be more satisfied with the information system if they believe that using the system will increase their performance and productivity (Adam Mahmood et al., 2000). Data content and system capability has a strong impact on perceived usefulness of the system and hence the designers have to carefully analyse the user requirements to determine their expectations and requirements of the content and accordingly incorporate relevant material and features into the system.

The data elements and structured of the STEP database is based on the requirements captured from the user needs assessment study. Participants were very contented with the fact that the database could provide them with different type of data (clinical, non clinical, in vitro and regulatory information) in one platform. Having the non-clinical and clinical studies under one platform would facilitate nonclinical to clinical correlations evaluations and integrated risk assessment (Monticello, 2015; Wallis, 2010). There are several ongoing efforts that focus on in-vivo toxicology and environmental chemicals (e.g. ActoR, Pubchem, Toxnet, echem portal). However none of these databases cover animal as well as human data. The National Library of Medicine's (NLM's) TOXNET database (Fonger et al., 2014) provides the human health effects and animal studies excerpts. It utilizes an extensive look-up-table of chemical structures and chemical names linked to either the central toxicity database websites or to specific records of primarily textual information pertaining to chemical toxicity. However, the user is constrained in the way in which they can search and access the data. The detailed characteristics of complete database (such as age, dose, route of administration or by types of articles, etc.) that is useful for computational purposes may be unavailable or difficult to obtain. This need has been satisfied by the STEP database as evident from the participants response to the information content. The response to the coverage of data elements in the STEP database validated that the major requirement of accommodating the clinical, non-clinical, in vitro and regulatory information in one place has been satisfied with the STEP database.

As the data in the STEP database comes from various internal and external sources, the data accuracy in terms of correctness and

Table 3
Breakdown on participants' success, problems, comments and changes implemented as per each task.

Functionality tested	Purpose of test	Tasks	Post task questions	Success	Problems	Possible reason for the problem faced	Changes implemented/ suggested
1. Log-in	To assess if the participant is able to register and login successfully in the database	To register to use the database	What page are you directed to when they were logged in the database?	Approximately 86 % participants were able to register and log in successfully and also provided right answer to the post task questions	7% found the registration process difficult and suggested making registration process more clear Participants commented that they had "difficulties figuring out how and where to register	The "registration form" for new user and "login form" for existing user was displayed on one page. It was speculated that this arrangement could have confused participants	Provided separate sections for registration and login in Version 2 of the database "Need help" link was provided with details on how to register in Version 2 Provided introduction on the first page in Version 2
			Rate the level of difficulty to complete the task	93% participants rated that it was easy to understand the registration process and to register and login in the database			
2. Search BY Excipients	To assess if the participants understands how to search for the information about a particular excipient in the database and to further refine the retrieved results	To perform the search based on excipients	Is there is any information on use of polysorbate 80 in children?	89 % participants were able to search for the excipient	Had no difficulty searching by the excipient	NA	NA
		To perform the search on excipient and filter the results to 'children' from age category column and 'topical' from route of administration column	Are there any studies reporting contact dermatitis in children on using the topical preparation containing polysorbate 80?			NA	NA
3: Results page—clinical data filtering	To assess if the participants understand what type of information is retrieved on results page and how the result page is structured or organized in different sections	To perform the search on a particular excipient and then find the information from general information section of the results page	If there is any Acceptable Daily Intake (ADI) information on sodium benzoate for paediatrics in the STEP database?	The performance success rate was good as 24 out of 29 participants successfully completed the task and answered the post task question correctly	NA	NA	NA
		To perform the search on a particular excipient and then find the information from clinical data section of the results page	Are there any studies reporting adverse effects of sodium benzoate if amounts are used above the ADI in paediatrics?	The performance success rate was very poor, only 12 out of 29 were able to perform the task successfully and answered the question correctly	Some found it very difficult to filter the columns to find the specific information while some indicated that it took a while for them to understand where the information would be and further to use the filter functionality to refine the data	The filter functionality provided in the database to sieve the data in the column was very similar to the filter option available in excel and thus needs knowledge of filtering columns in Excel. Some participants did acknowledge this. Some participants suggested providing the training on how to use the database specifically on how the information is organized in the database and how it	To address the difficulty of filtering the data, the video tutorials on "How the results page is structured and how to filter the data" was developed and is now available from the EuPFI website. Links to video tutorials and frequently asked questions are provided in Version 2

Table 3 (*Continued*)

Functionality tested	Purpose of test	Tasks	Post task questions	Success	Problems	Possible reason for the problem faced	Changes implemented/ suggested
						can be filtered to find specific data	
4: Results page—non clinical data Filtering	To assess if the participants were able to search for non-clinical information in the database	To filter the data specific to species i.e. dog and route i.e. oral	Find what effects are known for propylene glycol following an oral administration to dog?	27 out of 29 people successfully completed the task	NA	NA	NA
		To scan the adverse effects column for the NOEL information	what is No effect level (NOEL) for Propylene glycol in dog ?	21 participants were to able answer the second question correctly, 2 participants gave wrong answers while 6 participants were not able to find the information	Difficulty of filtering the data as seen for the task 3 above	Training/Help instructions on how to use the filters may be needed. It needs to be noted that though these tasks needed filtering the data as for task 3, high number of participants (72%) were able to perform the task as compared to task 3 (41%). The improvement in participant's performance to use the filters could be attributed to the participants getting used to the database and functionalities provided.	To address the difficulty of filtering the data, the video tutorials on "How the results page is structured and how to filter the data" was developed and is now available from the EuPFI website. Links to video tutorials and frequently asked questions are provided in Version 2
5: Search for excipients	To assess if the participants were able to perform complex queries and combine different attributes of clinical or non-clinical data to find more specific data	To perform a complex search with multiple fields	Which excipients are used intravenously in neonates? How many excipients did the participants get when searched for excipients used intravenously in neonates?	17 out 29 answered the post-task question correctly. 8 answered it incorrectly while 4 were not able to find the information	Incorrect selection of fields	The participant had to select 'preterm neonates, term neonates and paediatric age group'options from drop down menu to find the excipients used in neonates. There is possibility that the some participantss may have missed to select one of the options from 'preterm neonates, term neonates and paediatric age group'. The results indicated that that the participantss who gave incorrect answers have missed to select 'paediatric age group'	To avoid this confusion, it was decided to remove the 'paediatric age group' category and curate the information separately for individual age group on each row Changes are implemented in Version 2
6: Refine Specific Excipient using Search FOR Excipient	To assess if the participant understands how to use 'Search FOR Excipients' for performing very specific or complex queries	To limit the results at the search stage by refining the search, instead of limiting the results on the results page	To find carcinogenicity studies done in juvenile mice with benzyl alcohol	Approximately 86 percent of the participants completed first task successfully			To address this issue, in the revised version of the database, the ontology or list of the organs as per the system will be made available on STEP database webpage and a help icon will be provided next to organ/system category in the

Table 3 (*Continued*)

Functionality tested	Purpose of test	Tasks	Post task questions	Success	Problems	Possible reason for the problem faced	Changes implemented/ suggested
							database. On clicking the help icon it will list the organs/system ontology used for the STEP database. This will help participants to find the information as per the effect on organ/system
			To find the excipients that has studies reporting liver tumor in juvenile mice	65% of the participants were able to complete the second successfully	The major difficulty faced by the participants in performing the task was selection of the organ system. The difficulty experienced was in understanding the organ/system pathology/ontology	The database provides the upper-level nodes in the hierarchy of bodily systems such as circulatory system, the digestive system, and so on. It is difficult for some users to understand the ontology of organ system. Hence find it difficult to find the information for specific organ as per the body system	
7: Export the results	To assess if the participants are able to create a flexible, need based information reports that can be integrated it into different applications	To create a report in pdf format	Were you able to create the report and rate the level of difficulty	100% of the participants were able to perform the tasks successfully and found it very easy to export the searched and filtered data	NA	NA	NA

NA—not applicable.

consistency of data is very crucial requirement of the database. The standard curation procedure/manual (SCP) is developed and employed to improve the data accuracy. All curated data are captured in a structured manner to minimize the inconsistency amongst different curators. This is achieved by employing community-accepted controlled vocabularies and ontologies. Participants valued the curation efforts and acknowledged that this has helped them to retrieve the data efficiently. This validated that data accuracy requirement has been fulfilled.

Overall, the responses from the participants indicated that the STEP database provides vast and sufficient information that can be used to inform trade off decisions based on the current state of knowledge. Participants commented that major part of their work was done by the database by identifying the references relevant for

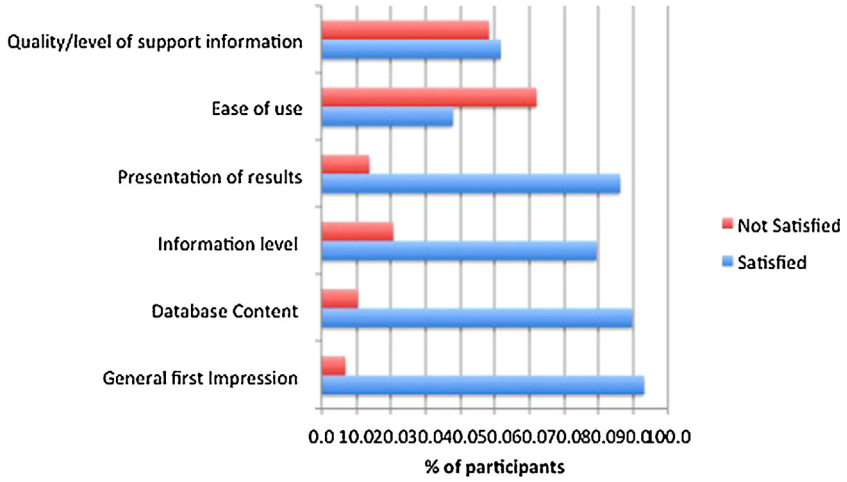

Fig. 4. User satisfaction.

Table 4
Likes and dislikes about the STEP database.

Database attributes	Participants comments on what they liked about the STEP database	Participants comments on what they didn't liked about the STEP database
Information content	Curated data: the curation/extraction of the information in table format made it easy to search for specific information (e.g. paediatric data) or refine the data for specific information (e.g. refine by route of administration)	Need more excipients
	The level of details	Need data on active ingredients
	Standardization of available data, multitude of references	Information on physicochemical characterization information of excipients
	Extensive information available	
	Mix of clinical and non clinical is very useful	
	It is very easy to use and very intuitive	
	Extraction to Excel/ PDF and filters are very useful	
Navigation and search interface	Simple structure and easy to use	Difficult to use the filters to limit the results for desired information
	Quick responses	Interface not immediately intuitive as a first time user but once familiar with database were able to perform the searches and refine results easily
	Easy to find regulatory status, ADI. Easy to filter for the required information.	User needs the knowledge of filtering the columns in excel to be able to use the filters comfortably
	Simple search by excipient—unfortunately limited number of excipients in the database	Not self-explanatory; difficult to understand the way the results are structured or took a while to understand where the information is
	Search function is relatively simple to use	It takes time to spot the nonclinical data tab and not obviously visible'
	Simplicity of use and specificity for paediatric use	Need more training and support information
	Ability to hide columns in the window and the way of searching	A training session on how to use the database would be helpful'

safety and toxicity of excipients. Also as the link to the original source is provided the users could always retrieve the full text as per their subscriptions availability.

5.2. System interface and ease of use

The results of this study confirm many of the findings of the earlier usability studies indicating that learn ability has a relatively smaller but significant effect on end-user satisfaction. The role of the designer is to facilitate the task for the user and to make sure that the user is able to make use of the system as intended and with a minimum effort to learn how to use it (Dillon, 2001). Despite the ratings on satisfaction being highest for information content and completeness, the low satisfaction with the ease of use and navigation was of concern in regards to the STEP database usability. The major limiting factor for the ease of use was the 'Filter' functionality that is provided to limit the data on results page. This is indeed the unique feature of the database and was introduced to facilitate the easy retrieval of the desired information by limiting the data on the results page, instead of having to perform another search again. However, users found it most difficult to use this feature. This was due to disappearance of the filter box when the mouse or cursor was moved out of the box. Every time the mouse was moved out of the 'filter box' the options selected by the user were reset and users had to select the options again. Though the users can be trained on how to use the filter or help option can be provided on page to demonstrate how to use the filter, it is not seen as an excuse to design a complex system. This is considered as one of design flaws and would be addressed as an area of significant concern and necessary improvement would be made to ensure the users would continue to use the feature and the database as intended.

There was some confusion caused by the phrases 'Search BY Excipients' and 'Search For Excipients'. This was not a major concern but few participants enquired on the differences between search by excipients and search for excipient. Some participants found it very self-explanatory that search by excipients allowed searches based on excipients and search for excipients allowed searches based on clinical data. Considering the specificity was needed to let users know exactly what the phrases meant, it was recommended to have one line explanation on the search page. No functions were identified as unnecessary.

Another significant finding in terms of learn ability was the need of user guidance. Training sessions and more support information was one of the common requests among the participants (40%). This finding is aligned with Lin et al. (1997) study that indicated a strong relationship exists between user guidance and learnability and usability of the software systems. Overall more help options and bolder eye-catching sidebars/ instructions and video tutorials would suffice the needs of the users in terms of ease of use or navigation. This would reduce the mental workload of the users since no extra efforts or time will be needed to carry out the desired task on the database. This would increase the productivity and efficiency of the user and thus improve the satisfaction and attitude towards the database.

Participants of the STEP database usability testing corroborated the finding that practically all user interfaces have learning curves that start out with the user being able to do nothing/limited activities at time zero when they first start using it to reaching a specified level of proficiency as they get used to the system (Nielsen, 1993).

5.3. Users and usability of the STEP database

User satisfaction seems to be driven by the users' motivation for visiting the database and cannot be reduced to a by-product of aesthetics, usability or even a combination of both. Participants paid enough attention to usability testing to point out numerous negative aspects, but they still acknowledged that this is just the first step towards integrating the vast information on safety and toxicity of excipients for paediatrics. It would thus appear that, while users are sensitive to usability issues and the focus on usability seems to be integral to the users experience, this need not affect their level of satisfaction with the interaction sufficiently to lower the satisfaction scores below a neutral level. Also it was important the users understand the purpose of the database. An

incorrect perception of the database might affect the user satisfaction. The detail information on purpose and scope of the database is available through EuPFI website. The target end-users includes the public, private sector, government agencies, hospitals and non-governmental organizations, particularly various sectors that have experience, expertise and demands of using excipients

6. Problem/changes prioritization

The study revealed several database design vulnerabilities and addressing those issues would enhance the usability of the database. To prioritize the results and determine which usability area or user needs were associated with greatest risks due to the observed usability problem, FMEA (Failure Mode Effect Analysis) model was adapted (Hertzum, 2006). Goddard (2000) stated that detailed software FMEA validates that the software has been constructed to achieve the specified requirements. It helped to recommend and prioritize changes or improvements to the STEP database design and functionalities. Each usability problem was treated as a failure mode. The problems that had impacted the important usability area and users task performance on users were identified. The problems with regards to number of participants affected were categorized in four groups; navigation, presentation, content and interaction. The metrics and dimensions of the FMEA matrix spreadsheet include: source (Usability scenario based test, Post test Questionnaire and Debriefing session); problem summary (one or two sentences describing problem); problem description (detailed explanation of problem); consequence (effect of problem on user); suggestion for improvement (one or two suggestions to mitigate problem); usability category (access, content, format, functionality, navigation, organization, symbols, terminology, workflow), database screen reference (screen/page of the database); severity of problem (1 = none, 3 = moderate, 5 = critical); probability of occurrence (1 = remote, 3 = moderate, 5 = very high); and risk priority (i.e. a product of severity and occurrence). The factors that contributed to the usability of the database and the satisfaction of the user needs include training, database awareness, end user confidence, productivity and efficiency.

A decision tree was developed (Fig. 5) to assess the severity of problem and accordingly scores were assigned. Severity of the usability problems is an important factor when defining the urgency of actions related to a problem (Hassenzahl, 2000). The tree provided a standard procedure for defining severity and helped to assign the severity in a consistent manner. It also provided the transparency to the stakeholders to help them make decision on changes to be made to the STEP database. Tables 5 and 6 lists the severity ranking and probability ranking.

The key objective was to use the FMEA "Recommended Actions" (Appendix B) field to rank and prioritize critical problems affecting the usability; develop and execute actions to eliminate the problems or reduce the effect on the usability of the STEP database. Fig. 6 represents the problems, severity and priority in which the problems will be solved.

Navigation and interaction were confirmed to be the critical problems that affected high number of users and have risk of user's not using the database if not improved. Hence the changes recommended to improve the navigation and interaction would be implemented on priority basis for next release of the database. While the problems related to additional data elements such as information on active pharmaceutical ingredient (API) was considered as minor and would be considered for future expansion of the database.

7. Implications for practice

The major output of the study is the "evaluation framework" for assessing the usability of the STEP database. The study provides a systematic method for the project manager to assess the usability of the database and identify the critical attributes and their impact on the ultimate usage of the database. The usability study might provide the long list of problems and recommended actions, but

Table 5
Problem probability ranking.

PROBABILITY of problem	Problem prob.	Ranking
Very high: problem is almost inevitable	>1 in 5	5
High: repeated problem	1 in 20	4
Moderate: occasional problem	1 in 50	3
Low: relatively few problems	1 in 80	2
Remote: problem is unlikely	<1 in 100	1

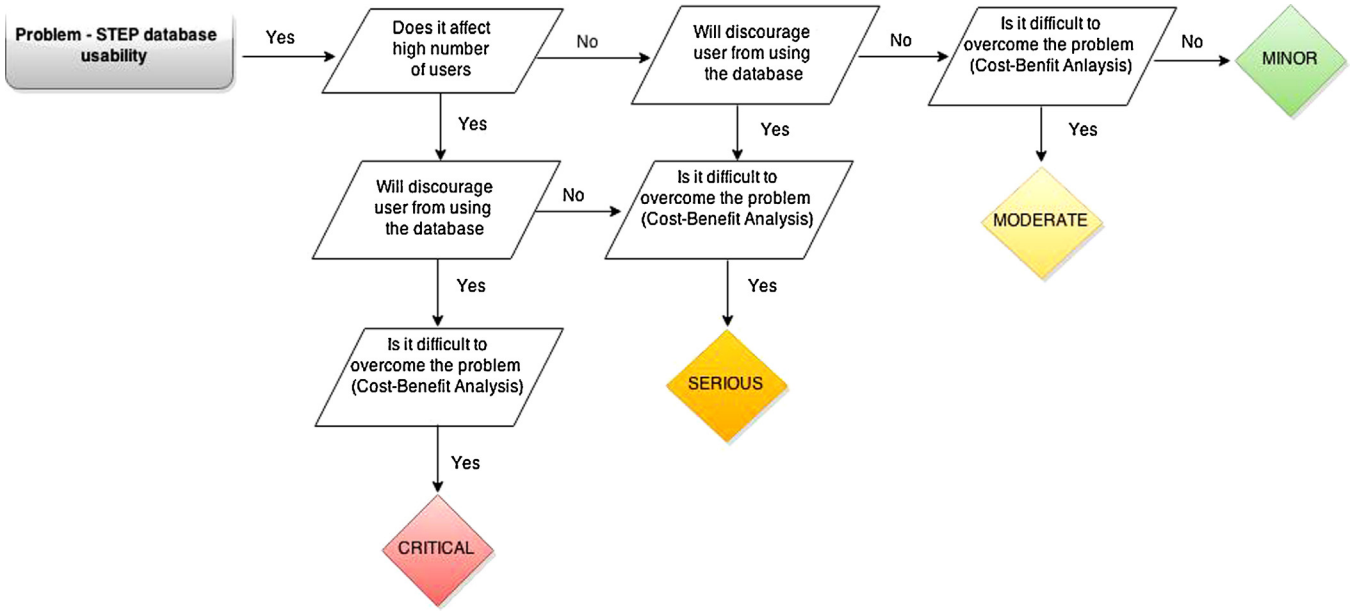

Fig. 5. Usability problem severity decision tree.

Table 6
Rating for severity of problem.

Problem	SEVERITY of problem	Ranking
Critical	Affects high number of users and the user could not continue without external help, or performed the task wrongly	5
Serious	Affects high number of users but the user continued after pausing for a significant time, or tried alternatives successful	4
Moderate	Does not affect high number of users and the user could not continue without external help, or performed the task wrongly	3
Minor	Does not affect high number of users and the user continued with a pause for thought	2
None	No effect	1

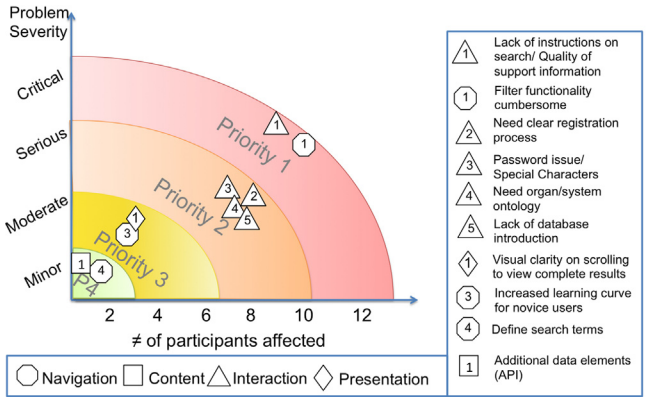

Fig. 6. Usability problem priority ranking.

the FMEA model used in this study will provide a mechanism to prioritize the preventive or corrective actions by assessing the relative severity and risk of the problem on usability of the database. It will act as a decision instrument for the stakeholders to prioritize the tasks that would be most likely to improve the usability of the database. The framework will help designers to refine the database as per user needs as the process provides clear understanding of the problems faced by the end-users and their requirements. For instance, the process can help developers understand the attributes (e.g. Navigation on search page, interaction on results page, content, presentation) that have the most impact on the user efficiency to use the database. The user suggestions, recommended actions and severity/risk priority number of the problem will help the developers to determine what kind of architectural changes can have the most meaningful impact. For instance, if user perceive the significant problems with the navigation of the database and has a high risk of putting the users off from the using the database, then the developers can focus efforts accordingly.

In summary, the evaluation framework has an analytical/problem-solving potential at any stage of database development or usability testing process.

8. Limitations

The users targeted in this study could not represent every type of user that would use this database such as medical information scientists, pharmacists, clinicians, care takers, poison information centres, etc. Thus, extending this to these users might help identify additional problems but not necessarily lead to different conclusions. The subset of users selected covered the key end-users of this database. Admittedly, however, this subset excluded health care professionals. The scenarios did not cover all possible ways the database can be searched but represented different points of impact focusing on core functionalities and the assignments from one professional to another, since these aspects were the main topics for the assessment. Issues such as user interface colours, buttons and minor functionalities were not as such part of the assessment. For task 5, the information could be retrieved by using different interface provided by the database (e.g. search by excipients and search for excipients). However, the way the question was asked did not clarify the method used by the user to find this information. This limitation was recognised at the stage of analysing the results. The lessons learned would be taken into consideration when designing the tasks and questions for subsequent testing of the database.

9. Conclusion

The lack of data on safety of excipients leads to uncertainties in paediatrics; so more information is urgently needed for optimal use of excipients in paediatric formulations. The STEP database is developed to facilitate rapid access to available toxicity data, a need shared by government, industry, consultants, and academic groups involved in drug development. Although the STEP database development is seen as a fairly mundane task, it acts as the foundation on which use and acceptability of excipients in paediatrics will be built and will fill the gap in excipients knowledge.

Developing a successful database requires the understanding the relationship between the data elements (user needs) and how efficiently the users can use the database. In many ways, this is the critical phase in that, having assembled the data in one place, building the linkages in a systematic manner will require understanding the value of the database and if the database stood up to the expectations of the users. Usability testing was performed as a part of development cycle of the STEP database to ensure it satisfies the need of the end-users and assessed the ease of use (navigation) and comprehension. Hybrid method of scenario based tasks and open-ended questionnaires helped assess a variety of questions ranging from analysis of specific interface problem to overall assessment of user expectations.

This study provides strong initial support for the usability of the STEP database. It is anticipated that the findings have addressed potentially problematic areas. The three critical attributes that were identified to have impact on the usability of the STEP database included (1) content and presentation of the results in the database (2) the navigation and search features (3) potential end-users of the database. Many problems encountered by the participants could be due to assuming the users would have the knowledge. Some elements were not laid out clearly and spelled out sufficiently for a novice's understanding. For instance, without clear outline of process of how to register and login, less computer savvy users were confused on the registration process. This study contributed to improving the database features in terms of enabling the non-expert users to use it more efficiently and purposefully. FMEA was adapted to prioritize changes or improvements to the STEP database design and functionalities. Refining the

database as per the recommendations will ensure that the database is more usable and accessible and thus will be used by more users and increase their participation towards the advancement of the database.

10. The journey continues

The STEP database should be seen as an open-ended process or journey—rather than a destination, in which the architectural changes are carried overtime and adapted to the changing unique needs of the end-user. The stakeholders, developers and maintainers are sensitive to the changing goals and objectives of the end-users of the STEP database. Hence it is important to understand the design and implementation issues and their ultimate effect on usage at any point with the database development lifecycle. In fact many of the users participated in the usability exercise to assess the effectiveness of the database and suggest changes accordingly. The project will follow an iterative user centered design-based evaluation approach, including assessment of the versions that will be later implemented, regarding the satisfaction of the requirements and usability criteria that were initially set. After implementation of changes to the STEP database, based on the usability evaluation, the evaluation would be repeated to determine how changes affect the system' usability. In this way the evaluation would be integrated in the process of further refinement and releases of the STEP database, iterative feeding information back into its continual improvement.

As the database is recently launched and new to the users, the issues identified are associated with novice users. Going forward as the database become more familiar within the users and as the learning curve is decreased, a study looking at expert users and novice users might identify the constraints and exceptions. It would be beneficial to perform longitudinal study that tests the proposed relationships as they unfold over time. Also the current objective of the usability testing was to assess how efficiently users are able to use the database and validate if it satisfied their needs in terms of content and ease of use. It does not assess how people will use the data within the database. To assess if the database is used appropriately, it will require understanding how scientific community, risk assessors etc. will use the database. The data presented in the database is 'as is' in the literature. The potential problem is that as this database is publicly available, the public may access the information and, without context, may misinterpret the information in the database. It is expected that users evaluate the relevancy and usefulness of the information with the context and with experts/toxicologists within their department. They are encouraged to refer to the original source. For future usability studies, it might be worthwhile to incorporate tasks/use cases to obtain some public feedback on how users may interpret the information in the database. This also can be achieved through annual EuPFI conference by bringing together the data generators, the data users and risk assessors. Future efforts will also focus on maintaining and updating data in the STEP database, and on expanding its capabilities by including more excipients.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007–2013 under grant agreement no. 261060.

The STEP database programme has assembled a tremendous intellectual resource ranging from EuPFI/USPFI consortium members involved in the programme to the information and data curation scientists at Synfo Data Services. They are the backbone of the database and their on-going commitment will assure that the program retains its focus. A special thanks to Brian Brown, (MLIS, Certified Usability Analyst, User Experience Librarian, National Institutes of Health Library) for his input and critical feedback in designing the usability questionnaire. Barbara Brandys (National Institutes of Health (NIH) Library, Division of Library Services, Office of Research Services, National Institutes of Health, Bethesda, MD, USA) is the USPFI content manager and has provided continual support to the STEP database project. She will lead the future usability testing of STEP database with US users to be carried in Q3 2015.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpharm.2015.06.016.

References

EMA, 2012. 5-year Report to the European Commission General report on the experience acquired as a result of the application of the Paediatric Regulation. http://ec.europa.eu/health/files/paediatrics/2012-09_pediatric_report-annex1-2_en.pdf (accessed 14.05.15.).

AAPS, 2014. 8th AAPS Italian University Netwrok (A.It.U.N.) Annual Meeting Medicines for Children's safe: challenges and opportunities, University of Pavia. http://www.aitun.it/en/events/8th-annual-meeting-pavia-march-2014.

Adam Mahmood, M.O., Burn, J.M., Gemoets, L.A., Jacquez, C., 2000. Variables affecting information technology end-user satisfaction: a meta-analysis of the empirical literature. Int. J. Hum.-Comput. Stud. 52, 751–771.

Calisir, F., Calisir, F., 2004. The relation of interface usability characteristics, perceived usefulness, and perceived ease of use to end-user satisfaction with enterprise resource planning (ERP) systems. Comput. Hum. Behav. 20, 505–515.

Dillon, A., 2001. Usability evaluation. In: Karwowski, W. (Ed.), Encyclopedia of Human Factors and Ergonomics. Taylor and Francis, London.

Faulkner, L., 2003. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav. Res. Methods Inst. Comput. 35, 379–383.

Fonger, G.C., Hakkinen, P., Jordan, S., Publicker, S., 2014. The National Library of Medicine's (NLM) Hazardous Substances Data Bank (HSDB): background, recent enhancements and future plans. Toxicology 325, 209–216.

Friedman, C., Wyatt, J., Owens, D., 2006. Evaluation and technology assessment. In: Shortliffe, E., Cimino, J. (Eds.), Biomedical Informatics. Springer, New York, pp. 403–443.

Goddard, P.L., 2000. Software FMEA techniques, Reliability and Maintainability Symposium, 2000. Proceedings. Annual, pp. 118–123.

Hassenzahl, M., 2000. Prioritizing usability problems: data-driven and judgement-driven severity estimates. Behav. Inf. Technol. 19, 29–42.

Hertzum, M., 2006. Problem prioritization in usability evaluation: from severity assessments toward impact on design. Int. J. Hum.-Comput. Interact. 21, 125–146. doi:http://dx.doi.org/10.1207/s15327590ijhc2102_2.

Jaspers, M.W., 2009. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int. J. Med. Inf. 78, 340–353.

Kaikkonen, A., Kallio, T., Kekalainen, A., Kankainen, A., Cankar, A., 2005. Usability testing of mobile applications: a comparison between laboratory and field testing. J. Usability Stud. 1, 4–16.

Kushniruk, A.W., Patel, V.L., 2004. Cognitive and usability engineering methods for the evaluation of clinical information systems. J. Biomed. Inf. 37, 56–76.

Levi, M., Conrad, F., 1998. Usability Testing of World Wide Web Sites, Research Papers of Bureau of Labor Statistics. http://www.bls.gov/ore/htm_papers/st960150.htm-SB5%20Return.

Lin, H.X., Choong, Y.-Y., Salvendy, G., 1997. A proposed index of usability: a method for comparing the relative usability of different software systems. Behav. Inf. Technol. 16, 267–277.

Monticello, T.M., 2015. Drug development and nonclinical to clinical translational databases: past and current efforts. Toxicol. Pathol. 43, 57–61.

Nielsen, J., 1993. Usability engineering, San Francisco, CA: Morgan Kaufmann. Academic Press.

Pope, C., Mays, N., 1995. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. Br. Med. J. 311, 42–45.

Power, C., Petrie, H.L., Mitchell, R.J., 2009. A Framework for Remote User Evaluation of Accessibility and Usability of Websites. Springer-Verlag, Berlin.

Powles-Glover, N.S., Edwards, J.M., 2011. The Lhasa Vitic Excipient (vehicle) Database: an information source on vehicle tolerability in rabbits. Reprod. Toxicol. 32, 172.

Qiu, Y., Yu, P., Hyland, P., 2007. A multi-method approach to assessing health information systems end users' training needs. Stud. Health Technol. Inf. 129, 1352–1356.

Quijano Ruiz, B., Desfontaine, E., Arenas-López, S., Wang, S., 2014. Pediatric formulation issues identified in paediatric investigation plans. Expert Rev. Clin. Pharm. 7, 25–30.

Root, R.W., Draper, S., 1983. Questionnaires as a software evaluation tool. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Boston, Massachusetts, USA, pp. 83–87.
Salunke, S., Giacoia, G., Tuleu, C., 2012. The STEP (Safety and Toxicity of Excipients for Paediatrics) database. Part 1—a need assessment study. Int. J. Pharm. 435, 101–111.
Salunke, S., Brandys, B., Giacoia, G., Tuleu, C., 2013. The STEP (Safety and Toxicity of Excipients for Paediatrics) database: part 2—the pilot version. Int. J. Pharm. 457, 310–322.
Tullis, T.S., 1998. A method for evaluating web page design concepts. CHI 98C Conference Summary on Human Factors in Computing Systems, ACM, Los Angeles, California, USA, pp. 323–324.
Virzi, R.A., 1992. Refining the test phase of usability evaluation: how many subjects is enough? Hum. Factors: J. Hum. Factors Ergon. Soc. 34, 457–468.
Vrazalic, L., 2003. Website usability in context: an activity theory based usability testing method. In: Whymark, G. (Ed.), Transformational Tools for 21st Century Minds. Knowledge Creation Press Pty., Ltd., Eveleigh, NSW, pp. 41–47.
Wallis, R.M., 2010. Integrated risk assessment and predictive value to humans of non-clinical repolarization assays. Br. J. Pharmacol. 159, 115–121.
West, R., Lehman, K., 2006. Automated summative usability studies: an empirical evaluation. Proceedings of the ACM SIGCHI 2006 Conference on Human Factors in Computing Systems, ACM Press. Montreal, April 2006 http://testrockit.azurewebsites.net/ (accessed April, 2015).
Yen, P.Y., Bakken, S., 2012. Review of health information technology usability study methodologies. J. Am. Med. Inf. Assoc. 19, 413–422.